

DOCUMENT RESUME

ED 194 622

TM 800 753

AUTHOR Parramore, Barbara M.: And Others
TITLE Effects of Mandated Competency Testing in North Carolina: The Class of 1980.
PUB DATE 20 Nov 80
NOTE 21p.: Paper presented at the Annual Meeting of the Evaluation Research Society (Washington, DC, November 19-22, 1980).
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Discipline Problems; Grade 11; *Grade Point Average; High Schools; *High School Students; *Minimum Competency Testing; Reading Achievement; *Remedial Instruction; Secondary School Mathematics; State Programs; Testing Programs; *Test Interpretation
IDENTIFIERS *North Carolina

ABSTRACT

Findings are reported based on an overall assessment of the impact of the North Carolina mandated competency test on eleventh graders of the class of 1980. The first part of the study examines the extent to which students who failed the first test showed improvement on subsequent testing. The second part studied those students who failed the initial test and passing students with whom they could be compared. It was found that a sample of students who failed initially showed improvement on re-examination and that the gains on the first retest were apparently related to remedial programs on an individual basis or in small group instruction. However, scores on subsequent retests remained fairly stable, the certain groups of students, especially the handicapped, showed little improvement. It was concluded, based on school-to-school differences, that the content of remedial courses is more important than how or how much is taught. It was also determined that neither test failure or success nor involvement in remedial classes affected students' final grade average or discipline status within the school. (MH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Effects of Mandated Competency Testing in
North Carolina: The Class of 1980

Presented at the
Annual Meeting
of the

Evaluation Research Society

Washington, D. C.

November 20, 1980

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

PARRAMORE

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Barbara M. Parramore
Robert C. Serow
James J. Davies

School of Education, North
Carolina State University
at Raleigh

Effects of Mandated Competency Testing in
North Carolina: The Class of 1980

Introduction

Beginning with the class of 1980, eleventh-graders in North Carolina were required to pass both the mathematics and reading sections of a Competency Test as a graduation requirement. The initial test was given during the fall of 1978 and those who failed one or both parts of the test had opportunities to participate in remedial programs and to take the test again in the Spring of 1979 and in the fall and spring of the next academic year. By June, 1980 all but a small fraction of the class of 1980, about two percent exclusive of dropouts, had been successful in passing the Competency Test.

The purpose of this study was to determine the effects of the Competency Test, if any, on the students academic and social behavior. In addition to reporting the overall results of the students' performance on the Competency Test, the first part of the study examined the extent to which students who failed the initial test showed improvements on subsequent retesting and to identify various factors which help to account for such improvements. In the second part, the sample was limited to students who failed the initial test (fall of 1978) and passing students with whom they could reasonably be compared. Here, the research questions dealt with 1) the associations, if any, between student performance on the Competency Test and changes in their class grade averages and the occurrence of reported behavior with discipline penalty imposed and 2) the association of the kind of remediation experiences with changes in grade average and disciplined behavior from before to after the testing.

Background

As part of the accountability movement affecting education during the 1970s, the General Assembly in North Carolina passed a law on June 9, 1977 requiring graduates to have demonstrated "those skills and that knowledge necessary to function independently and successfully in assuming the responsibilities of citizenship." The law also provided for remedial programs: "(s)tudents who fail to attain the required minimum standard for graduation in the eleventh grade shall be given remedial instruction and additional opportunities to take the test up to and including the last month of the twelfth grade." (N. C. House Bill 205, Chapter 522) Funding for the first year of the program was \$7.75 million dollars and during the second year, 1979-80, it was \$14.75 million with the increase reflecting the costs of remediation. Funds for remedial programs allocated to local education agencies depended upon the number of students failing the test and were contingent upon the local district accepting state guidelines for remedial programs.

The tests used to determine student competency were SHARP (reading) and TOPICS (math). Both are published by CTB/McGraw Hill. Members of the Competency Testing Commission, appointed by the Governor, were responsible for preparing the test specifications and selecting the tests. (For information on the test selection procedures, see Gallagher, 1980.)

Procedures

The findings reported here resulted from an overall assessment of the impact of the Competency Test on the class of 1980, the first group to participate in the North Carolina program. The original sample included 1,731 members of the class of 1980, drawn from ten high schools in four

local education agencies. In two districts, the entire cohort was included, in the other two, one or more high schools were selected. The sample consisted of 64 percent white and 36 percent black students. A small number of students who listed other or no racial identification were excluded. The students were from predominantly rural areas and lived in the eastern and central parts of North Carolina. The sample appeared to be broadly representative of the statewide test-taking population.

Data were gathered on students in the sample for three semesters, fall of 1978 and spring, 1979 (i.e., eleventh grade) and fall, 1979 (twelfth grade). Several types of information were collected:

Student background data (handicap, race, sex, and parents' education) were secured from records maintained by the State Department of Public Instruction.

Students' school records were compiled by staff members in the cooperating schools. Data were gathered on each of the following characteristics for each of the three semesters:

Average grade (all subjects) A through F,
Attendance (number of days absent),
Discipline (number of incidents reported to the principal's office
and penalties imposed), and
Remediation (average weekly amount and type of remediation--large
group, small group, individualized, combination, none--
for reading and math).

Competency Test scores (for the initial test and first and second retakes) were provided by the State Department of Public Instruction, using identification numbers rather than names in order to assure compliance with federal school privacy laws. These numbers were used by the researchers throughout the data collection and analysis.

Part I

Analysis and Results:

Table 1 shows the overall pattern of Competency Test performance within the original sample. Among whites, 91.5 percent of the girls and 88.9 percent of the boys, passed the Test on their first attempt, while corresponding figures for black girls and black boys was 58.6 and 50.8 percent, respectively. The total passing rate on the first Test was 72.5 percent. Upon completion of the second retest, these figures had increased to 97.8 percent of white girls, 95.5 percent of white boys, 77.8 percent of black girls and 71.0 percent of black boys. Overall, 6.8 percent of the students failing the initial testing continued to fail one or both parts of the Test and 7.7 percent did not take the Test again. (In North Carolina, the Competency Test may be taken as many as ten times, if necessary, up to the age of twenty-one.)

Table 1. Competency Test performances by race and sex*

Students	<u>Percent Passing</u>		<u>Percent Not Passing</u>	
	Initial Test	On Retests**	Took Retests	Dropout/No Retest
White girls	91.5	6.3	.6	1.6
White boys	88.9	6.6	2.3	2.2
Black girls	58.6	19.2	10.8	11.4
Black boys	50.8	20.2	13.4	15.6
Total	72.5	13.0	6.8	7.7

N = 1,731

* Includes results of initial testing and two retests

** Percent passing both math and reading but requiring make-up on at least one section of the test

Data on the extent and correlates of competency score change are shown in tables, 2, 3, and 4. Table 2 shows the test-to-test changes in

the raw scores (number items answered correctly) of students who failed one or both sections of the initial testing. (Students were required to retake only the section which had been failed.) Each section of the Test consisted of 120 items. To pass the math section, 77 correct answers were required whereas 87 were required to pass the reading section. Improvement in both reading and math was more evident on the first retest (spring, 1979) than on the second, administered during the fall of 1979. In the first retesting, math gains were greater than reading. Whites and girls tended to show greater improvement in both reading and math than blacks and males although these differences are statistically significant only on the first math retest. Non-handicapped pupils gained more on both sections of the first retest than did the handicapped.

Table 2. Average change in raw score by race, sex, handicap and remediation

Students	Initial Test to First Retest		First Retest to Second Retest	
	Math (N)	Reading (N)	Math (N)	Reading (N)
All students	17.5 (264)	10.8 (180)	3.3 (74)	3.8 (66)
Black	16.1 (186)	10.4 (139)	2.6 (60)	4.7 (54)
White	20.9 (78)	12.5 (41)	6.1 (14)	- .25 (12)
Male	15.8 (142)	10.3 (116)	3.1 (42)	3.8 (44)
Female	19.5 (122)	11.7 (64)	3.5 (32)	3.8 (22)
Handicapped	9.6 (55)	8.25 (48)	4.3 (26)	2.2 (22)
Non-handicapped	19.6 (209)	11.8 (132)	2.7 (48)	4.5 (44)
<u>Remedial Type</u>				
No remediation	13.3 (46)	9.2 (42)	7.8 (12)	8.5 (12)
Individual	20.1 (91)	12.0 (49)	3.2 (29)	6.9 (24)
Small group	19.1 (56)	11.5 (36)	2.1 (14)	3.1 (17)
Large group	15.8 (71)	10.7 (53)	1.4 (19)	-4.0 (13)
<u>Length of Remediation</u>				
None	13.3 (46)	9.2 (42)	8.8 (12)	8.5 (12)
1/2 Hour to 2 hrs.	13.4 (19)	5.9 (18)	8.0 (2)	6.4 (11)
More than 2 to 4	19.5 (95)	13.1 (76)	.23 (21)	-1.4 (11)
More than 4 hrs.	18.4 (104)	10.5 (44)	3.3 (39)	2.9 (32)

*Brackets indicate statistically significant ($p < .05$) difference between change scores of groups at ends or points of each bracket.

Competency Test score changes shown in Table 2 suggest certain inferences about the effectiveness of remedial programs. The most beneficial remediation programs provided students between the initial testing and the first retest apparently were either individual or small group instruction. This was true for both reading and math although only the math differences are significant. Students receiving no remediation improved at least several points less than did each of the remedial groups. However, this pattern does not hold for the fall, 1979 term, as non-remedial students showed somewhat greater gains from the first to second retest than did the remediated groups. The small size of some remedial groups in the fall, 1979 testing hinders the interpretation of the data. Generally, it would be expected that greater performance gains would follow from more intensive remediation. However, these data offer only limited support for such a proposition. The greatest improvements in both subject areas in spring, 1979 (initial test to first retest) were found among students receiving two to four hours of remediation each week, closely followed by the most time intensive remediation group, four or more hours. Again, this pattern does not fit the fall, 1979 data. Here, those without remediation gained most, followed by those who received only an hour or two of extra instruction.

Tables 3 and 4 show more clearly the correlates of Competency Test score change. To examine relationships between the student background, behavioral, and remediation variables and changes in Competency Test scores from the initial test to the first retest, the Statistical Analysis System (SAS) General Linear Model (SAS Institute, 1979) was applied. (Note that changes from the first to the second retest were generally minor and were not subjected to further analysis.) Basically, the General Linear Model procedure performs an analysis of covariance

using independent variables which can be either continuous or categorical. Using this procedure, an analysis was performed on the variance in change scores associated with each of the series of background, behavioral, and remedial variables. Also, a school variable was included to account for whatever influence each of the schools may have had on competency change. Such school to school differences would seem to be conceptually distinct from the two individual-level remedial variables (amount and type of remediation). This variable may not explain the nature of the school's influence on performance change but it might provide some indication of the relative importance of school-to-school differences.

The analysis was begun by removing variance attributable to students' prior test scores. Next, four background measures were entered (handicap, sex, parents' education and race), followed by the remedial, behavioral and school variables. The order of entry for these measures was as follows:

- Remedial type (individual, small group, etc.),
- Amount of remediation (during the remedial semester),
- Days absent (during remedial semester),
- Disciplinary incidents reported (during remedial semester),
- Average classroom grade (during remedial semester),
- School.

Examination of preliminary findings suggested several potentially significant interreactions among variables. These were introduced in the following order:

- Sex * Remedial type
- Sex * Amount of remediation
- Race * Remedial type
- Race * Amount of remediation
- Remedial type * Amount of remediation

As shown in tables 3 and 4, there are only a few variables which account for significant proportions of both math and reading change scores. The main one was the school-to-school difference, which explains nearly 18 percent of the gains in math and 10 percent of those in reading. In contrast, the two remedial variables (type and amount) together contributed only five or six percent of performance improvements in both math and reading scores. In this analysis, remedial type and amount were entered prior to the school variable and are thus assigned the variance which they share with the school. When this ordering was reversed, in a separate analysis, the effects of school-to-school differences became slightly stronger, while those of remedial type and amount dropped below the .01 level of statistical significance. The only other variables which showed a clear and consistent relationship with competency gains were students' handicap status (with non-handicapped students reporting greater improvement in both math and reading), and the interaction between sex and remedial type. Girls benefited from individualized instruction, while small groups were somewhat more effective for boys.

Table 3. General linear analysis of math score changes*

Source	Df	% Additional Variance		F	P
		Accounted For			
Initial math score	1	1.8%	7.23	.008	
Handicap	1	9.0	36.73	.0001	
Sex	1	1.5	6.09	.01	
Parent education	3	.5	.63	.60	
Race	1	1.7	6.88	.009	
Type of remediation	4	3.9	5.35	.001	
Average hours of remediation	2	1.1	2.21	.11	
Days absent	1	.9	3.92	.05	
Number of disciplinary incidents	1	.3	1.35	.24	
Average grade	1	1.8	7.25	.007	
School	10	17.7	7.28	.0001	
Sex * Type of remediation	3	2.1	2.88	.03	
Sex * Average hours of remediation	2	.1	.30	.74	
Race * Type of remediation	2	1.0	2.01	.13	
Race * Average hours of remediation	3	1.3	1.85	.13	
Type of remediation * Average hours	3	.5	.70	.55	
Total		45.2%			

* Change from initial test to first retest results

Table 4. General linear analysis of reading score changes*

Source	DF	% Additional Variance		F	P
		Accounted For			
Initial reading score	1	.0	0		.99
Handicap	1	2.7	6.06		.01
Sex	1	.5	1.11		.29
Parent education	3	1.7	1.23		.30
Race	1	.9	1.91		.17
Type of remediation	4	1.8	1.28		.28
Average hours of remediation	2	3.8	4.11		.02
Days absent	1	1.0	2.21		.14
Number of disciplinary incidents	1	.0	.02		.90
Average grade	1	.5	1.26		.26
School	10	10.0	2.16		.02
Sex * Type of remediation	3	4.0	2.86		.03
Sex * Average hours of remediation	2	4.7	3.35		.02
Race * Type of remediation	2	.1	.12		.89
Race * Average hours of remediation	3	1.5	1.62		.20
Type of remediation * Average hours	3	1.1	.81		.49
Total		34.3%			

* Change from initial test to first retest results.

Part II

Analysis and Results:

The second analysis of the Competency Test program consisted of all eleventh-graders from 11 high schools (in contrast to 10 high schools in the analysis reported earlier). The sample for this part of the study was limited to students who failed the initial test (fall of 1978) and passing students with whom they could reasonably be compared. The comparison group consisted of students who passed both the reading and math sections of the Competency Test by a margin of less than 10 points, and who, therefore, tended to be borderline in their performance on the Competency Test.

Two questions formed the basis for the focus of this phase of the study: 1) What are the associations, if any, between students' performance on the Competency Test and changes in their grade averages and in the incidence of disciplined behavior from before the first Test to two semesters after it? 2) Does remediation experience and the size of the remediation group have any association with changes in student grade average and disciplined behavior from before to after the testing?

An analysis of covariance was performed to determine whether or not significant variance in the change in grade average and disciplined behavior (behavior reported to the principal's office and penalty imposed) could be accounted for by student characteristics of race, sex, school, grade average and discipline status or by the interaction between these variables and performance on the test or type of remediation. The results of these analyses showed that race, sex, and school could not account for significant variance. However, the grade average and discipline status prior to the Competency Test did account for significant variance in the

change in grades and disciplined behavior.

Further analysis showed that students with average grades below C responded differently to remediation than did students with average grades of C or better. Since the grade of C is generally recognized as the division between satisfactory and unsatisfactory work, and since this division facilitates the presentation of interaction, grade average was treated as a nominal variable with students in categories of C or higher and below C grades. Behavior is also treated as a nominal variable with values of disciplined behavior and no disciplined behavior, depending upon the incidence of a student being reported to the administration for disciplinary reasons. With these divisions, four categories of students were established. These were:

1. Students with C or higher grade average and no disciplined behavior in the semester prior to the Competency Test.
2. Students with below C grade average and no disciplined behavior.
3. Students with C or higher grade average and disciplined behavior.
4. Students with below C grade average and disciplined behavior.

These students' performances on the Competency Test were then divided into four classifications. These were as follows:

1. Pass - students passing both the reading and math sections.
2. High fail - students failing only one of the two sections of the Test.
3. Moderately failed - students failing both sections but neither by more than 15 points.
4. Low fail - students failing one or both sections by more than 15 points.

Table 5 shows the extent of subsequent academic and behavioral change among students who failed or barely passed the initial Competency Test.

Three general points should be mentioned. First, there is some evidence of regression towards the mean with respect to grade change. That is, C or better students showed a slight decline, while gains were registered by those with below C averages. Second, while those with an initially favorable disciplinary status generally maintained that status, no clear pattern emerged for those with initially unfavorable statuses. Finally, with the consistent exception of those pupils with the lowest scores on the first Competency Test, most initial-failers managed to pass one of the re-examinations.

Table 5. Grade and disciplined behavior changes of failing and low passing score students on Competency Test

Student Classification	N	Mean grade average change 5-point scale	DISCIPLINE BEHAVIOR CHANGES				PERFORMANCE %	
			% with discipline behavior before but not after Test	% with discipline behavior before and after Test	% without discipline behavior before but after Test	% without discipline behavior be- fore and after Test	Took & Failed failed to keep both taking retests failed Test	
<u>C or better average/no disciplined behavior</u>	165	-.14					12.0	6.8
Pass	77	-.12			6	94		
High fail	47	-.19			19	81	5.7	7.5
Mod. fail	13	-.25			7	92	6.2	6.2
Low fail	28	-.12			7	93	50.0	21.0
<u>Below C /no disciplined behavior</u>	186	+.28					18.9	10.2
Pass	55	+.32			10	90		
High fail	35	+.28			9	91	11.1	8.9
Mod. fail	25	+.36			12	88	6.7	10.0
Low fail	71	+.22			18	82	41.5	19.2
<u>C or higher/ disciplined behavior</u>	22	0	81	18			11.6	12.5
Pass	14	0	72	28				
High fail	4	*	*	*				*
Mod. fail	1	*	*	*				*
Low fail	3	*	*	*				*
<u>Below C/ disciplined behavior</u>	46	+.15	38	61			26.3	23.7
Passed	11	+.14	64	36				15.8
High fail	12	+.18	66	33			5.3	15.8
Mod. fail	3		*	*				
Low fail	21	+.10	13	86			50.0	34.0

*Sample too small for reasonable analysis.

Table 6 offers a closer look at these relationships by breaking down the major change variables (grade average, incidence of disciplined behavior, and retest performance) according to the type of remediation to which students were exposed in the semester following the first Competency Test. Again, one hypothesis here is that certain remedial experiences, especially those offered in individualized or small group formats, will have generally positive effects on students' school performances. On the whole however, no such pattern is seen in Table 6, as changes in grades, discipline and competency levels seem to be distributed relatively evenly across all remedial groups.

Table 6. Mean grade change, disciplined behavior frequency and retest performance by remediation type for four grade/discipline categories of students failing the initial Competency Test

Student Classification	MEAN GRADE CHANGE (N = number)				% STUDENTS WITH DISCIPLINED BEHAVIOR IN 2nd SEMESTER AFTER TEST				% STUDENTS FAILING TO PASS THE TWO RETESTS				% STUDENTS FAILING TO TAKE BOTH RETESTS			
	<u>Remediation Type</u>				<u>Remediation Type</u>				<u>Remediation Type</u>				<u>Remediation Type</u>			
	None	Ind.	Sml.	Lrg.	None	Ind.	Sml.	Lrg.	None	Ind.	Sml.	Lrg.	None	Ind.	Sml.	Lrg.
	Groups				Groups				Groups				Groups			
C or higher average with no disciplined behavior	-.12 (91)	-.18 (36)	-.4 (11)	-.05 (26)	11.8	8.3	18.2	7.7	12	20	34	32	7.6	2.3	4.3	3.7
Below C average/no disciplined behavior	.33 (64)	.28 (50)	.32 (41)	.13 (31)	10.9	21.5	7.3	13.0	9	31	35	49	6.0	6.9	7.0	8.7
C or higher average*																
Below C average/ disciplined behavior	.11 (18)	.27 (11)	.25 (12)	-.05 (4)	55.0	81.2	46.0	25.0	48	43	50	75	27.0	7.0	16.0	25.0

* The number of C or higher average students with disciplined behavior who failed the initial test was too small for consideration.

Discussion

To return to the issues posed at the outset of this report, we have found that a sample of North Carolina students who initially failed the state Competency Test registered improvements on subsequent re-examinations, and that these gains were apparently related more to in-school factors than to individual background characteristics. However, these data raised several additional questions which might be considered in future studies of Minimum Competency Testing programs. First of all, it must be noted that while students posted impressive average gains on the first retest, scores on a subsequent examination tended to be fairly stable. In addition, certain groups of pupils (especially the handicapped) showed much less improvement than the majority of students. Another important issue concerns the pronounced differences in Reading and Math change scores. On the first retest, students gained an additional 17.5 correct answers in Math and 10.8 in Reading. Part of this discrepancy may be attributable to the likelihood that reading is more heavily determined by general sociocultural factors, and less susceptible to school influences, than are other academic skills (Coleman, 1975). Thus, in the present data set, the school and remedial variables accounted for nearly 23 percent of the variation in Math gains, but less than 16 percent in Reading. Unfortunately, the precise nature of this influence remains unclear. However, since school-to-school differences were found to be more important than gross methodological or quantitative measures of remediation, it may be reasonable to propose that what each school teaches in its remedial courses has greater impact on retest performance than how it teaches or how much is taught. As Coleman (1975) has suggested, the closer the linkage between test content and course content, the stronger will school effects appear to be.

The second major issue addressed in this report was the effects of students' participation in a Minimum Competency Test program or other types of school performances, notably discipline and class grades. There we found that neither Competency Test success or failure nor involvement in subsequent remediation seems to have much bearing on students' academic and social behaviors within the school. Accordingly, hopes or fears that Competency Testing programs will have broad spillover effects on high school students do not appear to be supported by these findings. However, in view of the importance attached to Competency Testing by educators, political leaders, and the public at large, it seems essential that research on these and other possible outcomes of Test programs continue to be pursued.

References

- Coleman, J. S. Methods and results in the IEA studies of effects of school on learning. Review of Educational Research, 1975, 45, 355 - 386.
- Gallagher, J. J. Setting educational standards for minimum competency: A case study. In R. Jaeger and C. Tittle (Eds.), Minimum competency achievement testing. Berkeley: McCutchan, 1980.